

ABSTRACT OF THE DISCLOSURE

A joint probability of two neighboring characters appearing in a given Japanese document database is statistically calculated. The calculated joint probabilities are stored in a table. An objective Japanese sentence is segmented into a plurality of words with reference to the calculated joint probabilities so that each division point of the objective Japanese sentence is present between two neighboring characters having a smaller joint probability.

10